In Vitro Permeation Test Studies for Topical Drug Products Submitted in ANDAs

Guidance for Industry

DRAFT GUIDANCE

This guidance document is being distributed for comment purposes only.

Comments and suggestions regarding this draft document should be submitted within 60 days of publication in the *Federal Register* of the notice announcing the availability of the draft guidance. Submit electronic comments to https://www.regulations.gov. Submit written comments to the Dockets Management Staff (HFA-305), Food and Drug Administration, 5630 Fishers Lane, Rm. 1061, Rockville, MD 20852. All comments should be identified with the docket number listed in the notice of availability that publishes in the *Federal Register*.

For questions regarding this draft document, contact (CDER) Susan Levine at 240-402-7936.

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)

October 2022
Generic Drugs

In Vitro Permeation Test Studies for Topical Drug Products Submitted in ANDAs

Guidance for Industry

Additional copies are available from:
Office of Communications, Division of Drug Information
Center for Drug Evaluation and Research
Food and Drug Administration
10001 New Hampshire Ave., Hillandale Bldg., 4th Floor
Silver Spring, MD 20993-0002
Phone: 855-543-3784 or 301-796-3400; Fax: 301-431-6353
Email: druginfo@fda.hhs.gov
https://www.fda.gov/drugs/guidance-compliance-regulatory-information/guidances-drugs

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)

October 2022
Generic Drugs

Contains Nonbinding Recommendations

Draft — Not for Implementation

TABLE OF CONTENTS

In Vitro Permeation Test Studies for Topical Drug Products Submitted in ANDAs Guidance for Industry[1]

This draft guidance, when finalized, will represent the current thinking of the Food and Drug Administration (FDA or Agency) on this topic. It does not establish any rights for any person and is not binding on FDA or the public. You can use an alternative approach if it satisfies the requirements of the applicable statutes and regulations. To discuss an alternative approach, contact the FDA staff responsible for this guidance as listed on the title page.

I. INTRODUCTION

This guidance is intended to assist applicants who are submitting abbreviated new drug applications (ANDAs) for liquid-based and/or other semisolid products applied to the skin, including integumentary and mucosal (e.g., vaginal) membranes, which are hereinafter called "topical products."[2] Because of the complex route of delivery associated with these products, which are typically locally acting, and the potential complexity of certain formulations, topical products (other than topical solutions) are classified as complex products.[3] This guidance provides recommendations for in vitro permeation test (IVPT) studies comparing a proposed generic (test) topical product and its reference standard (RS) for the purpose of supporting a demonstration of bioequivalence (BE) to the reference listed drug (RLD). The reference standard ordinarily is the RLD.[4]

[1] This guidance has been prepared by the Office of Generic Drugs in the Center for Drug Evaluation and Research at the Food and Drug Administration.

[2] Topical products in ANDAs within the scope of this guidance include ointments, creams, lotions, emulsions, pastes, shampoos, gels, suspensions, sprays, aerosols, foams, solutions and other semisolid and/or liquid-based dosage forms dispensed with a structured arrangement of matter (which may include more than one phase state).

[3] A *complex product*, as defined in the GDUFA Reauthorization Performance Goals and Program Enhancements Fiscal Years 2023–2027 (GDUFA III Commitment Letter) (accessible at https://www.fda.gov/media/153631/download, includes, among others, products with complex formulations (e.g., colloids) and complex routes of delivery (e.g., locally acting drugs such as dermatological products).

[4] A reference listed drug "is the listed drug identified by FDA as the drug product upon which an applicant relies in seeking approval of its ANDA." 21 CFR 314.3(b). A reference standard, which is selected by FDA, is the specific drug product that the ANDA applicant must use in conducting any in vivo bioequivalence testing required to support approval of its ANDA. See § 314.3(b). We recommend that the reference standard also be used for in vitro testing. There may be circumstances (e.g., when the RLD is no longer marketed) in which the reference standard is a drug product other than the RLD. For more information on RLD and reference standard products, see the guidance for industry *Referencing Approved Drug Products in ANDA Submissions* (October 2020). We update guidances periodically. For the most recent version of a guidance, check the FDA guidance web page at https://www.fda.gov/regulatory-information/search-fda-guidance-documents

This guidance does not address drug products that are administered via ophthalmic, otic, nasal, inhalation, oral, or injection-based routes, or that are transdermal or topical delivery systems (including products known as patches, topical patches, or extended release films).

It is beyond the scope of this guidance to discuss specific topical products to which this guidance applies. FDA recommends that applicants consult this guidance and any relevant product-specific guidances (PSGs)[5] and any other relevant guidances for industry,[6] when considering the design and conduct of IVPT studies that, in conjunction with other studies, as deemed necessary, may be appropriate to support a demonstration that a proposed generic topical product and its RLD are bioequivalent. FDA also recommends that applicants routinely refer to FDA's guidance web pages, because additional guidances may become available that could assist in the development of a generic topical product.

In general, FDA's guidance documents do not establish legally enforceable responsibilities. Instead, guidances describe the Agency's current thinking on a topic and should be viewed only as recommendations, unless specific regulatory or statutory requirements are cited. The use of the word *should* in Agency guidance means that something is suggested or recommended, but not required.

II. BACKGROUND

This guidance has been developed as part of FDA's "Drug Competition Action Plan,"[7] which, in coordination with the Generic Drug User Fee Amendments (GDUFA)[8] program and other FDA activities, is intended to increase competition in the market place for prescription drugs, facilitate the entry of high-quality and affordable generic drugs, and improve public health.

The Federal Food, Drug, and Cosmetic Act (FD&C Act) generally requires an ANDA to contain, among other things, information to show that the proposed generic drug product 1) is the same as the RLD with respect to the active ingredient(s), conditions of use, route of administration, dosage form, strength, and labeling (with certain permissible differences) and 2) is bioequivalent

[5] Generic drug product-specific guidances are available at FDA's Product-Specific Guidances for Generic Drug Development web page at https://www.fda.gov/drugs/guidances-drugs/product-specific-guidances-generic-drug-development.

[6] Other relevant guidances include the draft guidances for industry: *In Vitro Release Test Studies for Topical Drug Products Submitted in ANDAs* (October 2022) and *Physicochemical and Structural (Q3) Characterization of Topical Drug Products Submitted in ANDAs* (October 2022). When final, these guidances will represent the FDA's current thinking on these topics.

[7] See FDA Drug Competition Action Plan (describing the FDA's Drug Competition Action Plan, implemented in 2017 and designed to, among other things, further encourage robust and timely market competition for generic drugs), available at https://www.fda.gov/drugs/guidance-compliance-regulatory-information/fda-drug-competition-action-plan.

[8] In this guidance, *GDUFA* refers to the generic drug user fee program codified in the Generic Drug User Fee Amendments of 2012, Title III, Food and Drug Administration Safety and Innovation Act (Public Law 112-144), the Generic Drug User Fee Amendments of 2017, Title III, FDA Reauthorization Act of 2017 (Public Law 115-52), and the Generic Drug User Fee Amendments of 2022, Title III of Division F (the FDA User Fee Reauthorization Act of 2022) of the Continuing Appropriations and Ukraine Supplemental Appropriations Act, 2023 (Public Law 117-180).

to the RLD.[9] Thus, an ANDA will not be approved if the information submitted in the ANDA is insufficient to show that the test product is bioequivalent to the RLD.[10]

An IVPT study may be used to assess the rate and extent to which a drug (i.e., an active ingredient) from a topical product becomes available at or near a site of action in the skin, and may be used to characterize and compare the rate and extent of bioavailability for a drug from a test topical product and RS. The IVPT flux profiles resemble pharmacokinetic profiles and can be analyzed using unique IVPT endpoints that are somewhat analogous to the pharmacokinetic endpoints of maximum concentration (C_{max}) and the area under the concentration-time curve (AUC). Yet, IVPT studies characterize the rate and extent of absorption, not the distribution, metabolism and excretion that occurs in vivo. Therefore, while it is relevant to characterize the kinetics of topical drug bioavailability monitored by IVPT studies, the use in this guidance of the term "cutaneous pharmacokinetics" should not be construed to embody all aspects of pharmacokinetics—only those related to the absorption component that directly controls the rate and extent to which a topically applied drug becomes available locally at the site of action. This guidance focuses on general considerations and recommendations for the method development, method validation, and conduct of IVPT studies that are submitted in ANDAs and intended to support a demonstration of BE.[11]

III. IVPT METHOD DEVELOPMENT

The development of an IVPT method that is suitable to support a demonstration of BE for a specific topical product routinely involves a systematic series of exploratory studies. Inappropriate or insufficient efforts to develop an IVPT method that is suitable for its intended purpose increases the likelihood that the subsequent IVPT validation, pilot, and pivotal studies will ultimately be inadequate to support a demonstration of BE. By contrast, appropriate and systematic IVPT method development studies help to identify IVPT study designs and protocol (method) parameters which reliably produce flux profiles that can facilitate a comparison of the cutaneous pharmacokinetics of a drug delivered topically to the skin from test topical products and RSs.

A detailed and well-organized IVPT method development report should be submitted in an ANDA to show how the IVPT method was optimized, and to support a demonstration that the method parameters selected for the IVPT are appropriate or necessary, particularly in situations where the method parameters are different from the methods recommended in this guidance). The Agency's interest in reviewing the method development report is to understand why specific IVPT method parameters were selected and whether the resulting IVPT method is suitably sensitive and reproducible. This method development report should clearly indicate/distinguish

[9] See sections 505(j)(2)(A), (j)(2)(C), and (j)(4) of the FD&C Act (21 U.S.C. 355(j)(2)(A), (j)(2)(C), (j)(4)); see also 21 CFR 314.94.

[10] 21 CFR 314.127(a)(4), (6).

[11] A demonstration of no significant difference in the rate and extent of drug permeation into and through the skin of the test topical product and RS using an appropriately validated IVPT method can be used to support a demonstration of BE along with other data in the application (which may be specified in a PSG), as part of a comparative product characterization-based approach.

the method parameters used for each set of data, illustrate the efforts made to optimize the IVPT method, and demonstrate that the method parameters selected for the IVPT are appropriate.

Applicants are encouraged to use the recommendations in this guidance, and if an applicant elects to use methods that are different from those recommended in this guidance, the IVPT method development report should demonstrate why it is scientifically justified to use an alternative approach than what is recommended in this guidance to optimize the IVPT method.[12] Some examples of recommended procedures are described in subsequent sections, to help applicants identify circumstances when information should be submitted in the ANDA to explain why a different procedure was utilized.

A. IVPT Method Parameters

All relevant parameters of the final IVPT method should be summarized (e.g., in a table) and submitted in the ANDA. Also, information should be provided to briefly explain the choice of the final IVPT method parameters like the equipment (e.g., a vertical diffusion cell (VDC)), skin source (e.g., cadaver), skin type (e.g., posterior torso), skin preparation (e.g., dermatomed), skin barrier integrity test (e.g., trans-epidermal water loss (TEWL) measurement), skin barrier integrity test acceptance criteria (e.g., < 15 grams/meter^2/hour (g/m^2/hr)), topical product dose amount (e.g., 15 milligrams/centimeter^2 (mg/cm^2)), dose duration (e.g., 6 hours), study duration (e.g., 24 hours, 48 hours, etc.), receptor solution sampling times (e.g., 1, 2, 4, 6, 8, 12, 16, 20, and 24 hours), etc.

B. IVPT Method Considerations

The choice of some IVPT method parameters like the equipment, skin source, skin type, skin preparation, and skin barrier integrity test procedures may be based upon investigator experience or convenience, like the availability of specific equipment or instrumentation in a laboratory, established tissue supply agreements, or other logistical considerations. However, if the chosen IVPT method parameters do not appear to be well-suited for a specific IVPT method, it is the applicant's responsibility to systematically evaluate alternative method parameters, and ultimately, to validate that the IVPT method parameters chosen are suitable for the intended purpose. The recommended procedures for IVPT method validation are detailed in section IV of this guidance.

The choice of other IVPT method parameters like the topical product dose amount, dose duration, study duration (which may be longer than the dose duration), sampling schedule, sampling procedures, receptor solution composition, and sample analytical method may be different for each IVPT method, and such parameters of IVPT methods should be systematically developed, optimized, and/or validated for the relevant topical product, as appropriate. The IVPT method development studies should characterize how differences in these method parameters influence the resulting IVPT flux profile so that optimal study conditions can be objectively selected from among those evaluated.

[12] Applicants may choose to use an approach different from the approach recommended in this guidance. However, the alternative approach must comply with relevant statutes and regulations. See 21 CFR 10.115(d).

The selection of the dose amount used in the study should be assessed for each IVPT method based upon studies performed during IVPT method development. Different dose amounts may be compared in parallel on replicate skin sections from the same set of donors to optimize the dose amount for the IVPT study. Considerations for selecting an optimal dose amount may include (1) the consistency with which the dose can be applied (potentially using different dispensing and/or spreading techniques), (2) the reproducibility of the flux profiles, (3) the influence of dose amount and dose duration on the shape of the flux profile, and (4) the approximate range of drug concentrations in receptor solution samples at different time points (relative to the sample analytical method limits of quantification).

The selected sampling schedule and study duration should be sufficient to characterize the cutaneous pharmacokinetics of the drug, which ideally includes a sufficiently complete flux profile to identify the maximum (peak) flux and a decline in the flux thereafter across multiple subsequent time points. A dose that remains on the skin for the duration of the study may continue to deliver the drug for a sustained period and may not necessarily exhibit a suitable decline in the flux at later time points. In such instances, it may be appropriate to develop an IVPT method that involves wiping off the applied dose after a suitable duration on the skin and continuing to monitor the receptor solution for an extended period thereafter, during which the decline in the flux profile can be characterized. The sampling frequency should be selected to provide a suitable resolution for the flux profile, and a minimum of eight non-zero sampling time points is recommended across the study duration (e.g., 48 hours).

C. IVPT Method Procedures and Controls

Suitable technical procedures and control parameters should be established during method development. These may include procedures for preparing and mounting the skin on the diffusion cell in a consistent manner, determining the instrument settings that regulate the skin surface temperature within the specified range, performing the barrier integrity test appropriately, controlling the accuracy and precision of the dose amount dispensed on each skin section.

For example, a dosing procedure may be developed that uses a positive displacement pipette to dispense a volumetrically controlled amount of a topical product, targeting the deposition on the skin of a certain mass (e.g., 15 mg/cm^2) of topical product. If the inner diameter of the orifice in the dosing compartment of the diffusion cell is 15 millimeters (mm), and the effective dose area is ~1.77 cm^2, this would indicate a target dose of ~26.5 mg of topical product per diffusion cell. Experiments during method development may establish that, based upon the density of the topical product, a specific volumetric setting on a specific model of positive displacement pipette with a specific pipette tip repeatedly dispenses ~27.5 mg of topical product (e.g., characterized by multiple replicate pipette dispensations into a weigh boat on a fine balance). This pipette setting may be optimal for a dosing procedure where the dose spreading instrument, like the flat bottom of a high performance liquid chromatography (HPLC) glass vial, or the rounded end of a glass rod or capillary tube, is subsequently used to spread the dispensed dose evenly upon the skin section mounted in the diffusion cell, and where repeatedly weighing the dose-spreading instrument before and after the dose spreading indicates that the residual topical product remaining on the bottom of the glass vial after the dose spreading reproducibly amounts to ~1.0

mg of topical product (indicating that ~26.5 mg of the topical product would reproducibly be dosed to each skin section). Such characterizations of the technical procedures and control parameters for the IVPT method, like the reproducibility of the dosing procedure, should be established during method development and may not need to be demonstrated thereafter each time the same procedure is used.

D. IVPT Skin Barrier Integrity Testing: Common Methods

The technical procedures for the skin barrier integrity test should be established during IVPT method development. Three types of barrier integrity tests are common, however, there are currently no applicable compendial standard protocols or acceptance criteria for any of these three types of human skin barrier integrity tests. Nonetheless, recommended parameters for the three common types of barrier integrity tests are discussed below.

1. Trans-Epidermal Water Loss Skin Barrier Integrity Test

A TEWL skin barrier integrity test involves a measurement near the outer surface of the skin of the rate at which water (vapor) is fluxing through the skin barrier from the underside of the skin section. For the test, the skin section is mounted in a diffusion cell (e.g., clamped in place between the donor and receptor compartments), with the underside of the skin in contact with the receptor solution in the receptor compartment (e.g., phosphate buffered saline, pH 7.4), and equilibrated to a skin surface temperature of 32°C ± 1°C. If skin sections are cut large enough to cover the flange of the diffusion cell in which they are mounted, then after they have equilibrated for several hours at a skin surface temperature of 32°C ± 1°C, it may be feasible to gently remove the donor compartment without disrupting a skin section's adherence to the lower flange of the diffusion cell, thereby allowing the TEWL probe to be placed directly on the skin surface, instead of being placed atop the donor compartment. Typically, a minimum of three replicate measurements are made on each skin section, which are recorded after the measurements have stabilized.

Commercially available devices to measure TEWL may differ in design and operational principles. The TEWL measured by devices with certain designs (e.g., an open chamber versus a closed chamber) may be relatively more susceptible to the influence of environmental conditions. Therefore, environmental temperature and humidity are typically controlled as precisely as possible (e.g., a temperature range of 21°C ± 2°C and a humidity range of 50% ± 20% relative humidity are ideal, if feasible). More precise control of the relative humidity (e.g., in the range of 40% – 50%) may reduce the variability of TEWL measurements for devices with certain designs. Certain designs of measurement probes and adapters for in vitro use are available by the manufacturers of TEWL devices, and may be appropriate to use. Inconsistency in the diameters for the measurement probe chamber, the measurement probe orifice, the in vitro adapters, and the skin area being measured, as well as variation in the distance of the probe sensor(s) from the skin surface, potentially because of the (variable) height of donor compartments (when applicable), could increase the variability of TEWL measurements. Inconsistent control of the alignment of the TEWL measurement device in relation to the donor compartment and/or the skin section may also increase the variability of TEWL measurements. Also, the TEWL measured by devices with certain designs may be relatively more susceptible to

the influence of heat transfer from the hand that holds the probe. Applicants should follow relevant instructions in the manufacturer's user manual for the specific TEWL measurement device used.

No more than approximately 15 grams of water per square meter per hour (i.e., ≤ 15 g/m^2/hr) could be a reasonable skin barrier integrity acceptance (cutoff) criterion for a TEWL barrier integrity test on human torso or thigh skin; if this was selected as the cutoff criterion, skin sections with a TEWL > 15 g/m^2/hr would fail the test. Skin sections that fail a barrier integrity test should not be dosed, but may serve as non-dosed control skin sections. A higher cutoff (e.g., ≤ 20 g/m^2/hr) may also be reasonable if justified by experimental data demonstrating that the selected acceptance criterion appropriately discriminates skin sections with a compromised barrier integrity from those with a competent barrier integrity.

However, TEWL measurements for skin sections with a competent barrier integrity can vary depending upon the TEWL measurement device, the manner in which it is operated, and the environmental conditions (e.g., higher ambient humidity or greater distance from the skin surface may decrease the value of the TEWL measurement). Precise control of environmental and device/operational factors can minimize variability in TEWL measurements. Therefore, the technical procedures for measuring TEWL should be optimized during IVPT method development (or based upon prior optimization in the laboratory performing the test). Also, the TEWL measurement device should be appropriately calibrated (by the manufacturer, and for some devices, also before each set of tests). Applicants may provide information about the relevant calibration procedures specified by the manufacturer for the specific TEWL device used; this can be submitted in the ANDA along with the IVPT method development report, to support the appropriateness of the technical procedures established by the laboratory for TEWL measurements. When a TEWL barrier integrity test is used in any study phase (IVPT method development, pilot study, validation, and/or pivotal study) the ambient laboratory temperature and humidity during the TEWL barrier integrity test should be monitored and reported.

2. Tritiated Water Skin Barrier Integrity Test

An example of a recommended approach to a tritiated water skin barrier integrity test would be to mount the skin in a diffusion cell (e.g., clamped in place between the donor and receptor compartments) and allow it to equilibrate to a skin surface temperature of 32°C ± 1°C with the stratum corneum exposed to the air in the donor compartment and the underside of the skin in contact with the receptor solution (e.g., phosphate buffered saline, pH 7.4).

A small amount of tritiated water (sufficient to cover the entire surface of the skin section) would be briefly dosed on the stratum corneum. This dose of tritiated water would be left on the surface for a precisely controlled and relatively brief period (e.g., 5 minutes) after which it would be removed from the skin surface (e.g., using a pipette to remove the bulk volume and then an absorbent low lint laboratory tissue to gently blot dry). The receptor solution would then be sampled at a precise duration after the removal of the tritiated water from the skin surface (e.g., 30 minutes after the removal of the 5-minute dose of tritiated water from the skin surface).

While the entire volume of the receptor compartment may be removed and replenished, typically only an aliquot of the receptor solution (e.g., phosphate buffered saline, pH 7.4) is transferred to a suitable volume of scintillation fluid for counting. The volume of the aliquot typically depends upon the type of scintillation fluid used and the maximum amount of aqueous fluid that is suitable to mix with the scintillation fluid. A scintillation counter is then used to quantify the amount of radioactivity in the aliquot sampled, which can be used to calculate the amount of tritiated water that permeated into the larger (entire) volume of receptor solution; the calculation is performed using the specific activity of the tritiated water to equate a given amount of radioactivity to the equivalent volume of tritiated water that permeated per square centimeter of skin surface area.

Approximately 1.5 equivalent (eq.) microliter (µL) of tritiated water per cm^2 (i.e., ~1.5 eq. µL/cm^2 or ~1.5 eq. mg/cm^2) would be a reasonable skin barrier integrity acceptance (cutoff) criterion for a tritiated water barrier integrity test that involves a 5-minute dose followed by a 30-minute sampling duration (i.e., sampling 30 minutes after dose removal) on human torso or thigh skin. Skin sections with a tritiated water test result of > 1.5 eq. mg/cm^2 would fail the test and be excluded from the population of skin sections dosed with the topical product; skin sections that fail a barrier integrity test should not be dosed, but may serve as non-dosed control skin sections. Other acceptance criteria may also be reasonable if justified by experimental data demonstrating that the selected acceptance criterion appropriately discriminates skin sections with a compromised barrier integrity from those with a competent barrier integrity.

When calculating the results for a tritiated water barrier integrity test, it may be important to account for the surface area dosed. For example, if using an acceptance criterion of 1.5 eq. mg/cm^2 with a diffusion cell that has an orifice diameter of 15 mm and a skin surface area of 1.77 cm^2, the mass of tritiated water that would be calculated to have permeated into the receptor compartment would be ~2.7 eq. mg/cm^2 of tritiated water.

3. Electrical Based Skin Barrier Integrity Tests

There are several variations of electrical based skin barrier integrity tests that report the test result as a measure of the resistance, conductance, or a related electrical concept that characterizes the bulk flow of electrical current across the skin. Transepithelial electrical resistance tests involving the skin may be referred to more specifically as Trans-Epidermal Electrical Resistance (TEER) skin barrier integrity tests. The test results may be described in units of conductance, which is the reciprocal of resistance. Electrical based skin barrier integrity tests often use instruments that are designed to measure the inductance (L), capacitance (C), and resistance (R) of electronic circuits or electrical components; these instruments are commonly known as LCR meters and have different settings (test parameters) that can be adjusted.

An example of a recommended approach to a TEER skin barrier integrity test would be to mount the skin in a diffusion cell (e.g., clamped in place between the donor and receptor compartments) and allow it to equilibrate to a skin surface temperature of 32°C ± 1°C with the stratum corneum exposed to the air in the donor compartment and the underside of the skin in contact with an ionic solution (e.g., phosphate buffered saline, pH 7.4).

A small amount of the ionic solution (sufficient to cover the entire surface of the skin section) would be briefly dosed on the stratum corneum. Then, one lead/electrode from an LCR meter would be placed in contact with the solution in the receptor compartment while the other lead/electrode would be placed in contact with the solution in the donor compartment. After measuring the resistance across the skin (e.g., in kΩ, normalized for area, noting that resistance is inversely proportional to area) the solution in the donor compartment would be removed and the skin surface would be gently blotted dry with an absorbent low lint laboratory tissue. The skin (still mounted in the diffusion cell) would then be allowed to equilibrate with the dry air above for a sufficient duration to normalize the hydration state of the stratum corneum before being dosed with the test topical product or RS.

The results for a TEER skin barrier integrity test can vary substantially depending on the LCR meter settings (e.g., frequency) and the technical procedures used for the test. The acceptance criterion for a specific electrical based skin barrier integrity test method may be justified by experimental data demonstrating that the selected acceptance criterion appropriately discriminates skin sections with a compromised barrier integrity from those with a competent barrier integrity.

E. IVPT Skin Barrier Integrity Testing: General Considerations

There are three general considerations for the development or adoption of technical procedures for any skin barrier integrity test method during IVPT method development:

i. The technical procedures should not irreversibly alter the skin barrier. It may be acceptable to temporarily alter the hydration state of the stratum corneum by briefly depositing an aqueous solution on the surface of the skin, as long as sufficient time is afforded for the hydration of the stratum corneum to normalize before dosing of the topical product. The procedure described above for a brief (e.g., 5-minute) exposure of the skin surface to tritiated water followed by a 30-minute duration during which the hydration state of the stratum corneum is re-equilibrating would likely be appropriate. By contrast, a 30-minute exposure of the skin surface to an aqueous solution for an electrical-based test method, followed within 5 minutes by dosing of the topical product, may not be appropriate without further characterization of the influence of the hydration state of the stratum corneum on the discrimination sensitivity of the skin to differences in topical bioavailability. Similarly, if a portable lamp were placed close to the skin to improve visibility while study procedures were being performed, the heat from the lamp may alter the local (micro)environment of the skin in a manner that is not representative of the ambient environmental conditions in the laboratory; this should be avoided.

ii. The acceptance criterion should be a cutoff value for the test result, at which a skin section fails the test. Skin sections that fail a barrier integrity test should not be dosed but may serve as non-dosed control skin sections. Skin sections with a passing barrier integrity test result may be considered to have a competent barrier integrity and may be dosed. This acceptance criterion should be selected based upon an understanding of the distribution of test results (among multiple replicate skin sections from multiple donors) for the specific barrier integrity test procedure performed with the specific type and preparation of skin

under conditions relevant to the IVPT pivotal studies submitted in the ANDA. The intention of the barrier integrity test is to identify (and exclude) skin sections whose barrier integrity (intactness) is compromised. The intent is not to reduce the inherent variability in barrier function (permeability) in human skin that is representative of real variation in the human population. Also, the relative permeability of the skin to a drug from a topical product may not necessarily correlate with the permeability of the skin to water, and therefore, constraining the variability of the skin permeability to water (using a stricter acceptance criterion that excludes a larger number of skin sections) may not necessarily reduce the variability in the IVPT study results.

iii. The acceptance criterion should be able to discriminate skin sections with a compromised barrier integrity. This may be demonstrated by measuring the barrier integrity of skin sections mounted and equilibrated in a diffusion cell before and after deliberately compromising the skin barrier (e.g., by repeatedly using adhesive tape to strip away increasing amounts of the stratum corneum, piercing the skin several times with a 30 gauge needle, or using other physical or chemical insults to damage the skin barrier). Based upon the acceptance criterion selected, the test result for skin sections that pass the test before being damaged should fail the test after the damage.

F. Differences Between IVPT Method Development and Validation

1. Optimization of an IVPT Method Prior to Advancing to IVPT Method Validation

Different study designs and method parameters may be evaluated during the IVPT method development phase. For example, if the selected study parameters initially involve a dose duration of 48 hours and a study duration of 48 hours, and the flux profile is measurable, but it is not feasible to identify the maximum (peak) flux and a decline in the flux thereafter across multiple subsequent time points, then it may be appropriate to evaluate other study parameters as part of the IVPT method development. For example, a different target dose of the topical product and/or a longer sampling duration may be evaluated. An alternate study design may involve a shorter dose duration (e.g., 4–6 hours) after which the applied dose is removed from the skin, and the receptor solution continues to be sampled across a study duration that is sufficient to identify the maximum (peak) flux and a decline in the flux thereafter across multiple subsequent time points. While shorter dose durations can help to improve the shape of IVPT flux profiles, the removal of the topical product dose from the skin surface can be challenging and often requires its own method development and optimization. Also, the design of sensitivity studies for such an IVPT study design may require a more sophisticated understanding of IVPT studies. While reasonable efforts should be made to develop an IVPT method that produces a well-defined maximum (peak) flux and a decline in the flux thereafter across multiple subsequent time points, this may not be feasible for certain topical products even with study durations of 96 hours, or, at least, may not be feasible to produce reliably in all donors. In such circumstances, the IVPT method development report should detail the systematic efforts made to optimize the IVPT method.

2. Use of a Validated Sample Analytical Method for IVPT Method Validation

The IVPT method development studies, being exploratory in nature, are often performed using a sample analytical method that is not validated (e.g., an HPLC or ultrahigh performance liquid chromatography (UPLC) method, often involving mass spectrometry (MS)); also, IVPT method development studies are often conducted in a manner that is not compatible with a quality management system which would otherwise make the evidence generated suitable to support valid conclusions. Such method development studies would not be suitable to demonstrate the validity of an IVPT method, or associated results. Therefore, although it may appear to be redundant, certain experiments performed during IVRT method development may need to be repeated during IVPT method validation, using appropriate controls, like a validated analytical method and procedures that are compatible with a suitable quality management system.

It is important to clearly segregate and consistently identify those experiments and results that were part of IVPT method development separately from those that were part of IVPT method validation. It is also important to consistently identify all relevant method parameters and experimental conditions/controls for each set of IVPT results. Information in the method development report should clearly identify/distinguish when the results for apparently similar sets of experiments may have been obtained using different method parameters. Method development reports should clarify which sets of diffusion cells were run in parallel or separately (e.g., on separate days). In addition, the sample analytical method parameters used to analyze the samples from each set of IVPT experiments should be specified, and the report should indicate whether or not the sample analytical method was validated (either at the time of sample analysis or subsequently).

IV. IVPT METHOD VALIDATION

When all the relevant parameters of the IVPT method have been established, a pilot study should be performed using the final IVPT method and using a validated sample analytical method. The purpose of the pilot study is to validate the suitability of the selected IVPT method parameters by demonstrating that the performance characteristics of the IVPT method are appropriate to compare the cutaneous pharmacokinetics of a drug delivered topically from a test product and RS. The results from the pilot study, thereby, support the systematic validation of the IVPT method, which proceeds as a distinct study phase following IVPT method development.

The results from this IVPT pilot study can help to estimate the number of donors that may be needed to adequately power the IVPT pivotal study. In addition to the test topical product and RS evaluated in the pilot study, a parallel assessment should be performed with a third topical product or formulation that is known or designed to be different from the RS, to validate the selectivity of the IVPT method to discriminate differences in bioavailability. The IVPT pilot study results should be plotted with error bars, comparing the permeation profiles for the three treatment groups in the pilot study. Separate plots should be prepared for average flux results and average cumulative permeation results. These data can be used to support specific IVPT method validation parameters (e.g., permeation profile and range).

A pilot IVPT study performed with multiple skin donors (e.g., 4–6 skin donors) and a minimum of four replicate skin sections per donor per treatment group is recommended. As skin from an

increasing number of donors is evaluated in the pilot study, the accuracy of the estimated number of donors needed to adequately power the IVPT pivotal study may improve. While skin from the same donors evaluated in the pilot study may also be used in the IVPT pivotal study, the results from the pilot study should not be combined with the results from the IVPT pivotal study for the purpose of statistical analysis.

The equipment, methodologies, and study conditions used in the IVPT pilot study (and the eventual IVPT pivotal study) should be appropriately validated or qualified. If an applicant elects to use equipment, methodologies, or study conditions that are different from those recommended in this guidance, the applicant should demonstrate why it was necessary and scientifically justified to do so. Detailed protocols and well-controlled study procedures are recommended to ensure the precise control of dosing, sampling, and other IVPT study parameters, as well as potential sources of experimental bias.

The validation of the IVPT method should incorporate specific qualifications and controls (described below), performed using a validated sample analytical method, as applicable. The qualification of an IVPT method parameter refers to the process of defining what attributes make it suitable to perform its function in the IVPT method. For example, when repeated measurements of the temperature at the surface of skin mounted in a diffusion cell demonstrate that an IVPT equipment can maintain the skin surface temperature in the range of 32°C ± 1°C, the results can support a demonstration that the equipment is qualified to perform its function in an IVPT method for which a method parameter is the control of skin surface temperature in the range of 32°C ± 1°C across the relevant study duration.

A. Equipment Qualification

Suitable equipment for the IVPT method includes various models of VDCs and flow-through diffusion cells. The operating principles and specific test procedures differ among the various equipment; relevant procedures from the manufacturer may be used for installation, operational, and performance qualifications. The laboratory qualification of each diffusion cell should, at minimum, include 1) measurements of the diffusional area of the orifices of the donor and receptor compartments between which the skin is mounted, 2) the empirically measured volume of the receptor solution compartment in each VDC or the empirically measured outflow tube length for each flow-through diffusion cell, 3) the stability of the temperature measured at the skin surface (e.g., 32°C ± 1°C) across a relevant duration (e.g., 48 hours), and 4) the rate of stirring or agitation in VDCs, or the flow rate for flow-through diffusion cells, as applicable.

If information related to the diffusional area of the orifices and the volume of the receptor solution compartment for each diffusion cell is available from the manufacturer, that information should be provided for each relevant diffusion cell, in addition to the empirical measurements made by the laboratory performing the IVPT studies. The equipment should control the diffusion cell temperature so that the skin surface temperature is verified to be stable (e.g., 32°C ± 1°C) for each diffusion cell before dosing (e.g., measured by a calibrated infrared thermometer), and monitored periodically throughout the duration of the experiment by repeatedly measuring the skin surface temperature of a non-dosed control diffusion cell that is run in parallel with the other replicate dosed diffusion cells and connected to the same water bath or thermoregulation system.

B. Membrane (Skin) Qualification

Excised human skin is recommended as the membrane for the IVPT study. The validity of each skin section dosed in the study should be qualified using an appropriate test procedure to evaluate the stratum corneum barrier integrity. Acceptable barrier integrity tests may be based upon tritiated water permeation, TEWL, or electrical impedance/conductance measured across the skin. The test parameters and acceptance criteria used for the skin barrier integrity test should be justified for the specific method and instrumentation that is used during the study. The skin from all donors whose skin is included in the study should be prepared in a consistent manner and dermatomed to a relatively consistent thickness, within limits specified in the study protocol. The skin thickness should be measured and reported for each skin section included in the study. The assignment of replicate skin sections from a donor to each treatment group should be randomized, as feasible. It is acceptable to balance the distribution of skin thicknesses in each treatment group (test topical product or RS) by a procedure specified in the study protocol.

C. Receptor Solution Qualification

The composition and pH of the receptor solution used for the IVPT study should be qualified in relation to its compatibility with the skin as well as the stability and solubility of the drug in that receptor solution. The stability of the drug in the receptor solution samples should be validated as part of the receptor sample analytical method validation. The solubility of the drug in the IVPT receptor solution should be empirically determined in triplicate, to illustrate that the solubility of the drug in the receptor solution exceeds the highest sample concentration in the IVPT pivotal study, ideally by an order of magnitude. The solubility of the drug in the IVPT receptor solution should be sufficient to characterize the higher amounts of drug permeating from the increased drug delivery condition evaluated in the IVPT sensitivity assessment during IVPT method validation.

The inclusion of 0.1% polyoxyethylene[20]oleyl ether (also known as Oleth-20, Volpo-20, or Brij-20; CAS number 9004-98-2) is recommended to enhance the solubility of physiological buffer based (aqueous) receptor solutions for hydrophobic drugs. If additional solubility is needed, small increases in the concentration of polyoxyethylene[20]oleyl ether (e.g., from 0.1% or 0.2%, which is typically adequate for most hydrophobic drugs, to higher concentrations) are recommended, but should not exceed 6% polyoxyethylene[20]oleyl ether. Other strategies to improve the solubility of the drug in the receptor solution that may have the potential to alter the permeability of the skin (e.g., inclusion of organic solvents and alcohols in the receptor solution) are not recommended and may invalidate the IVPT method.

The inclusion of an anti-microbial agent in the receptor solution (e.g., ~0.1% sodium azide or ~ 0.01% gentamicin sulfate) is recommended to mitigate potential bacterial decomposition of the dermis and/or epidermis in the diffusion cell, regardless of the study duration. Other anti-microbial agents may also be acceptable, and if used, information should be included in the ANDA to explain the reason for their selection (and for the concentration at which they were used).

D. Receptor Solution Sampling Qualification

The accuracy and precision of receptor solution sample collection at each time point should be appropriately qualified. Evidence to qualify a sampling procedure should illustrate that the sampling technique can reliably collect a consistent volume of the sample from the well-mixed volume of the receptor compartment at each sampling event, and that no artifacts are likely to be created by the sampling technique. Information should be included describing the equipment manufacturer's specification for the accuracy and precision of receptor solution sampling, when available.

For IVPT studies using a flow-through diffusion cell, it may be appropriate to qualify the lengths of tubing, and their associated dead volumes, to accurately calculate the lag time before a sample elutes through the tubing and is collected. For IVPT studies using a VDC, removal of the entire receptor solution volume and full volume replacement of the receptor solution at each time point may provide optimal solubility sink conditions. The sampling of small aliquots of the receptor solution for an IVPT study may introduce anomalous measurements of apparently negative flux in certain regions of the IVPT study and produce flux profiles that are difficult to interpret.

E. Environmental Control

Ambient laboratory temperature and humidity during the study should be monitored and reported. An environmentally controlled temperature range of 21°C ± 2°C is recommended, and a humidity range of 50% ± 20% relative humidity is recommended, if feasible.

F. Permeation Profile and Range

The flux profile and cumulative permeation profile for the IVPT pilot study should be plotted across a range of sampling times, which corresponds to the IVPT pivotal study duration. The calculation of flux and cumulative total permeation is discussed in more detail below. The results of the IVPT pilot study should validate that the selected study parameters are suitable to adequately characterize the permeation profile (the cutaneous pharmacokinetics) of the drug within the selected study duration (the range of sampling time points).

A sufficiently complete flux profile should be adequate to identify the maximum (peak) flux and a decline in the flux thereafter across multiple subsequent time points in the IVPT pilot study. The results of the IVPT pilot study should also validate that the sampling frequency provides suitable resolution to adequately characterize the permeation profile (particularly the flux profile).

G. Precision and Reproducibility

The flux and cumulative permeation results from the IVPT pilot study (and the eventual IVPT pivotal study) should be calculated, tabulated, and reported for each diffusion cell at each time point, with summary statistics to describe the intra-donor average, standard deviation, and percent coefficient of variation (%CV) among replicates, as well as the inter-donor average, standard error, and %CV. Complete results for all data values used in the calculations should be

reported in a clear and organized manner, to facilitate the reconstruction of the flux and cumulative permeation results. The design of the study should be detailed and clear, and data values should be clearly associated with specific donors, replicates, treatment groups, time points, etc.

H. Dose Depletion

The recovery of permeated drug in the receptor solution should be characterized in each diffusion cell as the cumulative total permeation of the drug in the receptor solution over the IVPT duration. This may be expressed as a percentage of the nominal amount of drug in the applied dose (which may be estimated based upon the nominal strength of the drug in the topical product and the approximate mass of topical product dosed on the skin).

For example, if 10 mg of a topical product containing 5% drug was dosed on the membrane, the amount of drug in the applied dose may be estimated to be 0.5 mg (or 500 µg). If a cumulative total of 10 µg of drug diffused into the receptor solution across a 48-hour duration of the IVPT, it would be possible to estimate that the 500 µg dose would have been depleted by approximately 10 µg, amounting to an approximately 2% dose depletion. The average percentage dose depletion may thereby be estimated (not accounting for skin content) and should be reported.

I. Discrimination Sensitivity and Selectivity

The discrimination ability of the IVPT method may be described using two concepts: sensitivity and selectivity. The IVPT sensitivity studies are necessarily performed during IVPT method development to establish IVPT method parameters like the dose amount, dose duration, study duration, etc. However, the analysis of the results from these studies is qualitative in nature, and they need not be repeated during the IVPT method validation phase.

The IVPT sensitivity studies are typically performed toward the end of the IVPT method development phase, and a key purpose of these studies is to incorporate the final IVPT method parameters for the target dose and dose duration to be used in the pivotal study so that the IVPT sensitivity studies can support a demonstration of the validity of the final IVPT method. Therefore, IVPT sensitivity studies are described within this section of the guidance in the context of IVPT validation (rather than method development) to avoid dissociating the discussions of IVPT sensitivity (which is performed to establish the suitability of the final IVPT method parameters) and IVPT selectivity (which is performed once the final IVPT method parameters are established, and which is based upon the IVPT pilot study that is performed as part of the IVPT method validation). With the exception of the alternative dose amounts or dose durations used in the IVPT sensitivity study, it is important that the IVPT method parameters are consistent across the IVPT sensitivity, pilot, and pivotal studies (including the anatomical region specified in the study protocol (e.g., posterior torso), the skin source, and skin preparation).

1. IVPT Sensitivity

IVPT sensitivity is the ability of the IVPT method to detect changes in the cutaneous pharmacokinetics of the drug as a function of differences in drug delivery. If the IVPT method

consistently demonstrates higher and lower flux profiles (i.e., higher and lower values for IVPT endpoints) in response to increased and decreased drug delivery, respectively (or in response to other conditions expected to increase and decrease drug delivery, respectively), the IVPT method may be considered sensitive.

There are a few potential approaches by which to produce the differences in drug delivery that can be differentiated by a suitably discriminating IVPT method. Regardless of the approach used, the differences in the IVPT permeation profiles are not necessarily expected to be specifically proportional to differences in the dose amount, dose duration, or product strength. For example, three-fold differences in the dose amount (even if outside the recommended target dose range) may provide distinct flux curves but may not result in three-fold differences in the IVPT endpoints because the skin barrier may be rate-limiting both in vitro and in vivo.

In other words, if the target dose for the pivotal IVPT study was 10 mg/cm^2, a 3-fold lower dose would be ~3 mg/cm^2 and a 3-fold higher dose would be 30 mg/cm^2; thus, an IVPT sensitivity study might compare the flux profiles from 3, 10, and 30 mg/cm^2 doses of the topical product. Similarly, if the target dose for the pivotal IVPT study was 15 mg/cm^2, a 3-fold lower dose would be 5 mg/cm^2 and a 3-fold higher dose would be 45 mg/cm^2; thus, an IVPT sensitivity study might compare the flux profiles from 5, 15, and 45 mg/cm^2 doses of the topical product. An IVPT sensitivity study performed with multiple skin donors (e.g., 4–6 skin donors) and a minimum of four replicate skin sections per donor per treatment group is recommended.

- **Modulation of Dose Amount**: An IVPT method development study with different dose amounts may provide supportive evidence that the IVPT methodology is sensitive to differences in drug delivery.

  This approach is well suited to topical products that contain volatile components that evaporate from the formulation following dose application to the skin. Modulating the dose amount for such topical products effectively alters the thickness of the applied dose. The majority of volatile components from a thinner dose will tend to evaporate more rapidly (compared to a thicker dose), and a thinner dose will tend to deliver less drug into the skin (and/or for a shorter duration) compared to a thicker dose.

  Modulating the dose amount can be an effective technique to modulate differences in drug delivery for formulations with volatile components, like gels, lotions, and many creams. However, modulating the dose amount may not necessarily produce perceptible differences in drug delivery for topical products like petrolatum-based ointments, or other types of topical products that do not evaporate on the skin, or that may not experience dose-dependent differences in metamorphosis that can alter the rate and extent of drug delivery.

- **Modulation of Dose Duration**: For many topical products, it may be more effective to modulate the dose duration, instead of the dose amount, to produce differences in drug delivery and associated changes in the cutaneous pharmacokinetics of the drug.

An IVPT method development study with a controlled dose amount (e.g., 15 mg/cm^2) dosed for different durations (e.g., 2 hours, 6 hours, and 12 hours) may be well suited to provide supportive evidence that the IVPT methodology is sensitive to differences in drug delivery from many topical products. The scenario described in this example would support an IVPT study design where a topical product dose of 15 mg/cm^2 is dosed for 6 hours (the target duration for the IVPT study) and then wiped off. The applied dose may be removed with a series of cotton-tipped swabs, one or more of which may be dry and one or more of which may be moistened (e.g., with a soap solution or water). The initial (dry) swab typically removes the bulk of the dose and subsequent swabs are used to remove the residual dose (i.e., the residue of the topical product which may otherwise continue to deliver drug into the skin) and/or to rinse the skin.

To support a demonstration of the sensitivity of the IVPT study, the permeation profile produced by the target dose duration for the IVPT study (e.g., 6 hours) should be compared with a shorter dose duration (e.g., 2 hours) that is expected to perceptibly decrease the drug delivery, and also be compared with a longer dose duration (e.g., 12 hours) that is expected to perceptibly increase the drug delivery. Thereby, the three dose durations compared in the IVPT sensitivity study are designed to produce perceptible changes in the cutaneous pharmacokinetics of the drug as a function of differences in drug delivery, and thereby support a demonstration of the sensitivity of the IVPT method.

The specific dose durations may be selected based upon an initial exploratory IVPT study performed during IVPT method development that characterizes the permeation profile when the dose is left on the skin for a longer duration (e.g., 24 or 48 hours). An important feature of the results from such an IVPT study is the duration of the initial phase of the permeation profile, when the flux is increasing at a relatively rapid rate.

For example, if such an exploratory study indicates that the flux increases on a steep slope until approximately 12 hours, and then continues to deliver the drug at a gradually increasing rate thereafter, it may suggest that the permeation profile for a dose duration of longer than 12 hours (e.g., 24 hours) may not be perceptibly different from that of the 12-hour dose duration, especially when compared in a relatively small number of donors and replicates (e.g., four donors with four replicates each per dose duration). It may also suggest that a 12-hour dose duration may be a good choice for the longest of the three dose durations in the IVPT sensitivity study.

The target dose duration should be selected based upon considerations like the sensitivity of the sample analytical method, the ability to produce a permeation profile that can be perceptibly discriminated from that produced by the longer (12 hour) dose duration, and/or the labeled use of the topical product (which may indicate that the topical product should be reapplied every 4–6 hours).

The shortest of the three dose durations in the IVPT sensitivity study should be selected based upon the sensitivity of the sample analytical method and its ability to produce a permeation profile that can be perceptibly discriminated from that produced by the target (6 hour) dose duration.

- **Modulation of Product Strength**: To validate the sensitivity, specificity, and selectivity of an in vitro release test (IVRT) method, altered strength formulations are routinely prepared. While it may seem convenient to use these altered strength formulations in an attempt to demonstrate the sensitivity and selectivity of an IVPT method, doing so may not produce the desired outcomes. There may be circumstances when this strategy may produce perceptibly different permeation profiles, however, in many instances, the resulting permeation profiles may not be perceptibly different when compared in a relatively small number of donors and replicates (e.g., four donors with four replicates each per topical product strength). In general, the modulation of topical product strength to support a demonstration of IVPT sensitivity is not recommended because it may not consistently produce the expected increase or decrease in drug delivery; however, in certain situations, higher and lower strength formulations (relative to the nominal strength of the RS) may suitably increase and decrease the drug delivery and cutaneous pharmacokinetics relative to that from the nominal strength topical product.

2. IVPT Selectivity

IVPT selectivity is the ability of the IVPT method to discriminate the cutaneous pharmacokinetics of the drug between the RS and a topical product or formulation that exhibits differences in drug delivery relative to the RS. The IVPT pilot study with the parallel assessment of the RS, the test topical product, and a third topical product or formulation that is known or designed to be different from the RS may provide supportive evidence that the IVPT methodology is selective for differences in drug delivery. Topical product batch information for all topical product lots used in IVPT method development, validation and pilot studies, as applicable, should be submitted in the study reports. The topical product information should include, but not be limited to, information about the batch formula, manufacturing date, batch size, altered manufacturing processes (if applicable) and, if available, potency and content uniformity. The evaluation of inequivalence may be based upon a qualitative or quantitative comparison of the permeation profiles and/or the IVPT endpoints.

J. Robustness

A primary assumption related to robustness testing is that the test system performs consistently when all system variables (e.g., temperature, stirring rate) are at nominal settings. A value of robustness testing is that it can verify whether the system continues to provide a consistent output when specific variables are slightly altered, thereby qualifying operational ranges for those variables. However, the variability inherent in the permeability of human skin, whether in vitro or in vivo, may not be compatible with the primary assumption related to the consistency of the test system.

Nonetheless, results from studies during IVPT method development that appear to support the robustness of the IVPT method or system should be reported, if relevant. For example, an IVPT method may be robust to substantial variations in the stirring rate of the receptor compartment. Similarly, the permeation profile of a drug into and through human skin may appear to be robust to certain differences in the topical product strength. Ultimately, because it may not always be

feasible to validate the robustness of IVPT method parameters, IVPT study procedures should be controlled as precisely as possible.

V. SAMPLE ANALYTICAL METHOD VALIDATION

While exploratory studies performed during IVPT method development may use an unvalidated sample analytical method, it is essential that all studies conducted as part of the IVPT method validation use a validated sample analytical method. A validated IVPT method should use a validated sample analytical method (e.g., HPLC/MS or UPLC/MS). Therefore, a discussion of the sample analytical method for the IVPT method is included in this guidance under this section on IVPT method validation.

However, the study protocols and reports related to the IVPT method are distinct from those for the sample analytical method that is used to quantify drug concentrations in IVPT receptor solution samples. The validation of a sample analytical method, in and of itself, does not demonstrate the validity of an IVPT method. Separate and specific reports should be submitted for the validation of the sample analytical method (e.g., HPLC/MS or UPLC/MS) and for the validation of the IVPT method.

Any results from studies of the IVPT method that are performed during method development using a different sample analytical method than that which is ultimately validated, cannot support a demonstration of the validity of the IVPT method. Information should be provided in the IVPT method validation report referencing the (separate) sample analytical method validation, and clearly indicate that all relevant results in the IVPT method validation report were obtained using a validated sample analytical method (as opposed to a sample analytical method with different parameters than those which were validated).

The receptor sample analysis procedures (e.g., typically involving an HPLC/MS or UPLC/MS system) should be performed using chromatography software (e.g., a chromatography data system) with audit trails, and should include a multi-point (6–8 concentration) calibration curve with suitable quality control samples, and should be validated in a manner compatible with the FDA guidance for industry *Bioanalytical Method Validation* (May 2018).

The validation of the receptor sample analytical method should include relevant qualifications of dilution integrity, if applicable, as well as stability assessments with the highest relevant temperature in the receptor solution for the longest relevant duration; the highest relevant temperature may be warmer than 32°C because the temperature of the receptor solution is often higher than the temperature at the surface of the skin, and the longest relevant duration may be the longest interval between sampling time points for methods in which the entire receptor solution is replaced at each sampling time point, or it could be longer in scenarios with only partial sampling of the receptor solution (e.g., 34°C for 48 hours).

If the samples are processed in specific ways for analysis (e.g., by drying and reconstituting the receptor samples in a smaller volume to concentrate the sample and increase the effective analytical sensitivity, or by dilution of receptor solution samples into the validated curve range of

the sample analytical method) those procedures should be validated (e.g., by qualifying the dilution integrity during the sample analytical method validation). The stability of the drug in the receptor solution sample should be validated in a receptor solution matrix that has been exposed to the underside of the skin in a diffusion cell under conditions relevant to the IVPT pivotal study.

VI. IVPT PIVOTAL STUDY

The IVPT pivotal study protocol should incorporate considerations relevant to BE studies, in general.

A. Handling and Retention of Samples

Refer to 21 CFR 320.38, 320.63 and the FDA guidances for industry *Handling and Retention of BA and BE Testing Samples* (May 2004) and *Compliance Policy for the Quantity of Bioavailability and Bioequivalence Samples Retained Under 21 CFR 320.38(c)* (August 2020), as applicable, regarding considerations for retention of study drug samples and to 21 CFR 320.36 for requirements for maintenance of records of BE testing. Retention samples should be randomly selected from the drug supplies received before allocating topical product units for use in an IVPT study in which the test topical product and RS are compared.

B. Control of Study Procedures

Study procedures that have the potential to influence the results of the study should be appropriately controlled. Also, experimental observations that may have the potential to influence the interpretation of the study results, as well as any protocol or standard operating procedure (SOP) deviations, should be reported.

Control of procedures related to the skin include the consistent control across the study of the skin preparation (e.g., dermatoming of skin sections) and the thickness of skin sections mounted on diffusion cells, as well as the skin storage conditions, including the duration for which the skin was frozen and the number of freeze-thaw cycles to which the skin was exposed. Skin from the same anatomical location should be used from all donors, and the demographics (age, race, sex) should be reported for all donors. Also, the IVPT sensitivity, pilot, and pivotal studies should use skin from the same anatomical site; otherwise, if skin from different anatomical sites is used across the different study phases, it may not be possible for the results of the IVPT sensitivity and pilot studies to support a demonstration of the discrimination ability of the IVPT method used for the pivotal study because the method parameters would not be aligned across the respective studies. Similarly, if a non-rate-limiting support membrane is used beneath the skin section (e.g., a filter membrane used in a validated IVRT method for the same topical product) then it should be used in a consistent manner for the IVPT sensitivity, pilot, and pivotal studies.

Control of procedures related to the dose include the control of the area of dose application, the dose amount, the dosing technique, the dose duration, and the blinding and randomization

procedures for dosing. The test topical product and RS should be dosed in an identical and consistent manner for all diffusion cells in the study. Differences in dosing technique may alter the metamorphosis of the dosage form on the skin, and inconsistencies in the diameter of the area dosed on each diffusion cell may significantly influence the dosed area and contribute to errors in the calculation of flux.

Control of procedures related to sampling include the control of sampling time precision, the sampling technique, the duration of sampling and replacement of receptor solution, the sample volume or flow rate, and sample handling and storage.

Control of procedures related to the pivotal study should include a non-dosed control skin section from each skin donor, which should be mounted in a diffusion cell and otherwise treated identically to the dosed skin sections, including sampling of the receptor solution at all time points to ensure that drug concentrations monitored in the receptor solution are associated with the dose applied in the IVPT pivotal study, and not drug contamination in the skin from that donor that might permeate into the receptor solution across the duration of the study. A pre-dose "zero" sample collected from each diffusion cell is also recommended, which may identify potential contamination associated with each skin section and/or each diffusion cell.

In addition, investigators should perform the IVPT validation and pivotal studies within a quality management system that includes, but is not limited to, documented procedures for:

- Study personnel identification, training, qualification, and responsibilities
- Study management and study management personnel responsibilities
- Quality control (QC) and QC personnel responsibilities
- Quality assurance (QA) and QA personnel responsibilities
- Use of SOPs
- Use of study protocols
- Use of study reports
- Maintenance and control of the study facility environment and systems
- Qualification and calibration of instruments and computerized systems
- Good documentation practices including, but not limited to, contemporaneous documentation of study procedures and recording of experimental observations or deviations from procedures specified in the study protocol or in relevant SOPs
- Maintenance of suitable records that facilitate the reconstruction of study events and procedures, including study sample handling and storage records (e.g., sample tracking

logs), audit trails for sample analysis procedures, control of study materials and reagents, and electronic data control

- Archival of study records

C. Blinding Procedure

A detailed description of the blinding procedure should be provided in the study protocol and final report. The packaging of the test topical product and RS should be similar in appearance to maintain adequate blinding of the investigator and any experimental operators.

D. Randomization

The method of randomization should be described in the protocol of the IVPT study and the randomization schedule provided, preferably in a SAS data set in .xpt format (created using the SAS XPORT procedure). It is recommended that an independent third party generate and hold the randomization code throughout the conduct of the study to minimize bias. The applicant may generate the randomization code if not involved in the packaging and labeling of the test topical product and RS dosed in the study. A sealed copy of the randomization scheme should be retained at the study site and should be available to FDA investigators at the time of site inspection to allow for verification of the treatment identity of each skin section.

E. Dosing

In the IVPT pivotal study, the test topical product and RS should be dosed in an alternating pattern on successive diffusion cells (skin sections) from each donor. One of two dosing sequences (illustrated below) may be randomly assigned for each donor:

a. ABABAB…
b. BABABA…

F. Study Design

The IVPT pivotal study should compare the cutaneous pharmacokinetics of the drug from the test topical product versus that from the RS using excised human skin with a competent skin barrier mounted on a qualified diffusion cell system. The IVPT pivotal study should use a design that directly compares the test topical product and RS on skin from the same set of donors, each with the same number of replicate skin sections per donor per treatment group (dosed with either test topical product or RS topical), using the same IVPT method parameters.

The IVPT pivotal study design, methodology, and diffusion cell equipment considerations relating to sampling precision should be controlled as precisely as possible. For example, it may be appropriate to stagger the dose application on successive diffusion cells and to synchronize the sampling time points with the dosing time for that diffusion cell, to ensure consistent durations between dosing and sampling of all diffusion cells.

G. Inclusion Criteria

In general, the following inclusion criteria should apply: healthy, normal, barrier-competent skin from male and/or female donors of at least 18 years of age. Inclusion criteria related to donor demographics (e.g., age, race, sex) should be specified in the study protocol and demographic information should be reported for each donor. Additional criteria may be added by the applicant.

The skin may be harvested following excision from patients undergoing a surgical procedure or excised from cadavers. A consistent source is recommended for all the skin used. The anatomical region specified in the study protocol (e.g., posterior torso) should be consistent for all donors whose skin is included in the study.

The study protocol should specify the inclusion (acceptance) criteria for skin sections based upon the barrier integrity test result, which should be reported for each skin section.

The study protocol should specify inclusion criteria related to the temperature and duration of skin storage as well as the number of freeze-thaw cycles, all of which should be reported for each donor's skin.

The study protocol should specify the inclusion criteria related to the skin harvesting/processing procedures and skin thickness (e.g., dermatomed skin of 500 μm ± 250 μm thickness) used in the IVPT study.

H. Exclusion Criteria

In general, the following exclusion criteria should apply. Skin from subjects with a known (history of) dermatological disease should be excluded from the study. Skin with tattoos, stretch marks, or any visible sign of abnormality should be excluded from the study. Skin exhibiting a significant density of terminal hair is not recommended and should be excluded from the study. Additional criteria may be added by the applicant.

While gentle washing or rinsing of the skin surface is appropriate, submerging the skin in an aqueous solution for more than a few minutes may damage the skin barrier and should be avoided; such skin sections should be excluded from the study. Also, skin that has been subjected to shaving with a blade; abrasive polishing; tape-stripping; or cleansing with alcohols, solvents, or other strong solutions that could damage the skin barrier should be excluded from the study.

Skin from donors with significant background levels of the drug or other compounds that may interfere with the quantification of the drug in receptor solution samples should be excluded from the study.

Skin from donors exhibiting a high barrier integrity test failure rate among replicate skin sections may be excluded from the study, and skin from an alternative donor may be used instead.

I. IVPT Endpoints

The endpoints for the IVPT pivotal study are based upon parameters that characterize the rate and extent to which the drug permeates into and through the skin and becomes available in the receptor solution. Specifically, the rate of drug permeation is characterized by the flux (J) and the extent of drug permeation is characterized by the total cumulative amount (AMT) of drug permeated into the receptor solution across the study duration.

The flux (rate of drug permeation) should be plotted as J on the Y-axis in units of mass/area/time (e.g., nanograms (ng)/cm^2/hr) versus time on the X-axis. Flux profiles commonly resemble plasma pharmacokinetic profiles, however, it is important to distinguish that the flux is a rate, rather than a concentration. The extent of drug permeation should also be plotted, as the total cumulative amount (AMT) of drug permeated on the Y-axis in units of mass/area (e.g., ng/cm^2) versus time on the X-axis.

The flux should be calculated based upon: the receptor sample concentration (e.g., 2.0 ng/mL) at each time point; the precise, empirically measured volume of that specific diffusion cell (e.g., 6.0 mL) which may vary between individual cells; the area of dose application (e.g., 1 cm^2); and the duration for which the receptor volume was accepting the drug. For example, if the sample exemplified here represented a 2-hour period following dosing, then J would be calculated based upon the values above as:

$$J = [(2.0 \text{ ng/mL}) \times (6.0 \text{ mL})]/(1 \text{ cm}^2)/(2 \text{ hrs}) = 6 \text{ ng/cm}^2\text{/hr}$$

This flux should be calculated and reported for each diffusion cell for each sampling interval and plotted across the entire study duration to generate the flux profile for each diffusion cell. The rate calculated above may be plotted at the 2-hour time point, or at the midpoint between 0 and 2 hours (i.e., 1 hour).

In addition, the AMT should be calculated and reported for each diffusion cell. This cumulative amount of drug that has permeated (in total across the entire study) should be reported as the AMT endpoint, rather than using a trapezoid rule to calculate the area under the flux curve.

The maximum flux (J_{max}) at the peak of the drug flux profile and the AMT should both be compared for locally-acting test topical products and RSs. This is somewhat analogous to the comparison of the C_{max} and AUC for systemically-acting test products and RSs, inasmuch as the pair of endpoints in each case facilitates a comparison of the rate and extent to which the drug from each type of product (locally-acting or systemically-acting) becomes available at the site of action.

A confidence interval (CI) should be calculated for each IVPT endpoint:

a. the natural log-transformed maximum flux (J_{max})
b. the natural log-transformed total cumulative amount (AMT) permeated

It is the responsibility of the applicant to determine the number of donors required to adequately power the IVPT pivotal study, however, a minimum of four dosed replicates per donor per treatment group (test product or RS) is recommended.

At the completion of the study, if the number of skin replicates is the same for all donors in the test topical product and RS treatment groups in the IVPT study, a statistical analysis for a balanced design is recommended. If skin sections or diffusion cells are excluded from the final statistical analysis because of experimental loss/issues, and the resulting data set is unbalanced, a statistical analysis for an unbalanced design is recommended.

Approaches to statistical analysis of the pivotal study are described in section VIII of this guidance. Appendix I provides example SAS code for determining BE with both a balanced dataset and an unbalanced dataset. Appendix II provides numerical examples with simulated data sets. Appendix III provides example R code for determining BE.

VII. SUBMITTING INFORMATION ON IVPT STUDIES IN AN ANDA

For IVPT studies with topical products submitted in ANDAs that are intended to support a demonstration of BE, detailed study protocols, relevant SOPs, and detailed reports should be submitted for the IVPT method validation (including the IVPT pilot study) and the IVPT pivotal study. In addition, a detailed report describing the IVPT method development should be submitted. These protocols, SOPs, and reports should be submitted in module 5.3.1.2 of the electronic Common Technical Document (eCTD) and should describe experimental procedures, study controls, quality management procedures, and data analyses.

Note that the study protocols, SOPs, and reports related to the IVPT method are distinct from those for the sample analytical method that is used to quantify drug concentrations in IVPT receptor solution samples (e.g., an HPLC/MS or UPLC/MS method). Separate protocols and SOPs should be submitted for the sample analytical method validation. Sample analytical method development and validation reports, pilot and pivotal IVPT study sample analysis reports, as well as associated SOPs and protocols relevant to the sample analysis of an IVPT study with human skin should be submitted in Module 5.3.1.4 of the eCTD.

VIII. IVPT PIVOTAL STUDY STATISTICAL ANALYSIS

The two treatment groups would correspond to the test topical product (T) and the RS (R). The statistical analysis should consider a sample of n donors, for which r_j^T replicate skin sections from the j^{th} donor ($j = 1, \cdots, n$) are available for the T group and r_j^R replicate skin sections from the j^{th} donor ($j = 1, \cdots, n$) are available for the R group. Each replicate (i) from each donor (j) should have been randomly assigned to each product.

Define the following quantities:

- T_{ij} = the natural log-transformed IVPT endpoint (J_{max} or AMT) dosed with the test topical product for the i^{th} skin replicate from the j^{th} donor ($i = 1, 2, \cdots, r_j^T,\ j = 1, 2, \cdots, n$)
- R_{ij} = the natural log-transformed IVPT endpoint (J_{max} or AMT) dosed with the RS for the i^{th} skin replicate from the j^{th} donor ($i = 1, 2, \cdots, r_j^R,\ j = 1, 2, \cdots, n$)
- r_j^T = the number of skin replicates from the j^{th} donor dosed with the test topical product ($j = 1, 2, \cdots, n$)
- r_j^R = the number of skin replicates from the j^{th} donor dosed with the RS ($j = 1, 2, \cdots, n$)
- $r^* = r_1^R + r_2^R + \cdots + r_n^R$ = the total number of skin replicates in the R group
- n = the number of donors

If the numbers of skin replicates available for the final statistical analysis are the same for the n donors for the T group and the R group, the resulting data set is *balanced*. For simplicity of notation, the common number of skin replicates for one donor for one treatment group in a balanced data set is denoted as $r = r_1^T = r_2^T = \cdots = r_n^T = r_1^R = r_2^R = \cdots = r_n^R$.

A diffusion cell may be excluded from among the replicates in a data set when there is a documented observation of a failure (e.g., visual observation that a skin section tears and leaks during the study) or a protocol deviation (e.g., the receptor compartment in a diffusion cell is discovered to be empty at the first sampling time point). In such instances, if sufficient skin remains from the same donor, and no samples from that diffusion cell have been analyzed, a replacement diffusion cell can be set up and studied. Otherwise (if the diffusion cell cannot be replaced) the resulting data set becomes unbalanced.

The statistical analysis methods for assessing BE in the cases of a balanced data set and an unbalanced data set are described below. For a donor to be included in the statistical analysis, there should be at least 3 replicate skin sections from the donor for each (T and R) treatment group.

Step 1.

Determine S_{WR}, the estimated within-donor standard deviation of the RS, for each of the natural log-transformed IVPT endpoints J_{max} and AMT:

$$S_{WR} = \left(\frac{\sum_{j=1}^{n} \sum_{i=1}^{r_j^R} \left(R_{ij} - \bar{R}_{\cdot j} \right)^2}{r^* - n} \right)^{1/2}$$

where $\bar{R}_{\cdot j} = \frac{1}{r_j^R} \sum_{i=1}^{r_j^R} R_{ij}$ is the average of log-transformed observations across all r_j^R replicates from donor *j* dosed with the RS.

(a) If $S_{WR} \geq 0.294$, use the scaled average BE (SABE) approach to determine BE for the individual IVPT endpoint(s) in **Steps 2, 3.1, and 4.1**

(b) If $S_{WR} < 0.294$, use the regular average BE (ABE) approach through the two one-sided tests (TOST) procedure (Schuirmann, 1987) to determine BE for the individual IVPT endpoint(s) in **Steps 2, 3.2, and 4.2**

Step 2.

Determine the point estimate for the mean difference of T and R products ($\hat{I}$), its standard error ($se(\hat{I})$), and the corresponding degrees of freedom (df^*).

For a **balanced** data set, determine $\hat{I}$, $se(\hat{I})$, and df^* by the following:

- $\hat{I} = \bar{I}_{\cdot} = \frac{1}{n}\sum_{j=1}^{n} I_j$ where $I_j = \frac{1}{r}\sum_{i=1}^{r}(T_{ij} - R_{ij})$
- $S_I^2 = \frac{1}{(n-1)}\sum_{j=1}^{n}(I_j - \bar{I}_{\cdot})^2$ (estimate of inter-donor variability)
- $se(\hat{I}) = \sqrt{S_I^2/n}$
- $df^* = n - 1$

For an **unbalanced** data set, approximate $\hat{I}$, $se(\hat{I})$, and df^* by using PROC MIXED (or PROC GLM) in SAS. The example code is provided in Appendix I.

Step 3.1. Scaled Average BE (SABE) Approach

In the SABE approach, the hypotheses to be tested are:

$$H_0: \frac{(\mu_T - \mu_R)^2}{\sigma_{WR}^2} \geq \theta$$

$$H_a: \frac{(\mu_T - \mu_R)^2}{\sigma_{WR}^2} < \theta$$

where:

- $\mu_T - \mu_R$ = mean difference of T and R products
- σ_{WR}^2 = within-donor variance of R product
- $\theta = \frac{(\ln(m))^2}{(\sigma_{W0})^2}$, $m = 1.2500$ (BE limit), and $\sigma_{W0} = 0.25$ (regulatory constant)

Rejection of the null hypothesis supports the conclusion of equivalence of the two products.

Determine (1-α)*100% upper confidence bound for $(\mu_T - \mu_R)^2 - \theta\sigma_{WR}^2$ based on Howe's Approximation (Howe, 1974) (α = 0.05):

$$X + Y + sign(V) * |V|^{1/2}$$

where:

- $X = \hat{I}^2 - se(\hat{I})^2$
- $Y = -\theta S_{WR}^2$
- $X' = \left(|\hat{I}| + t_{(1-\alpha),df^*} * se(\hat{I})\right)^2$
- $Y' = -\theta \frac{(r^*-n)S_{WR}^2}{\chi^2_{(1-\alpha),(r^*-n)}}$
- $V = \left(X' - X\right) * |X' - X| + \left(Y' - Y\right) * |Y' - Y|$
- $sign(V) = 1$ if $V > 0$; 0 *if* $V = 0$; -1 if $V < 0$

Note that $t_{(1-\alpha),df^*}$ is $(1-\alpha) * 100^{\text{th}}$ percentile of the Student's t-distribution with df^* degrees of freedom and $\chi^2_{(1-\alpha),(r^*-n)}$ is $(1-\alpha) * 100^{\text{th}}$ percentile of the Chi-square distribution with $(r^* - n)$ degrees of freedom.

Step 3.2. Regular Average BE (ABE) Approach

In the ABE approach, the hypotheses to be tested are:

$$H_0: \mu_T - \mu_R \leq -\ln(m) \text{ or } \mu_T - \mu_R \geq \ln(m)$$
$$H_a: -\ln(m) < \mu_T - \mu_R < \ln(m)$$

where:

- $\mu_T - \mu_R$ = mean difference of T and R products
- $m = 1.2500$ (BE limit)
- $\ln(m)$ is the natural logarithm of the BE limit

Rejection of the null hypothesis supports the conclusion of equivalence of the two products.

Determine the (1 − 2α)*100% confidence interval for $\mu_T - \mu_R$ (α = 0.05):

$$\hat{I} \pm t_{(1-\alpha),df^*} * se(\hat{I})$$

where $t_{(1-\alpha),df^*}$ is $(1-\alpha)*100^{\text{th}}$ percentile of the Student's t-distribution with df^* degrees of freedom.

Step 4.1. BE Determination with SABE Approach

For the test product to be bioequivalent to the reference standard, **both** of the following conditions must be satisfied for each IVPT endpoint tested:

a. the 95% upper confidence bound for $(\mu_T-\mu_R)^2-\theta\sigma_{WR}^2$ must be less than or equal to zero (numbers should be kept to a minimum of four significant figures for comparison).

b. the point estimate of the T/R geometric mean ratio must fall within the pre-specified limits $\left[\frac{1}{m},\ m\right]$, where m = 1.2500.

Step 4.2. BE Determination with ABE Approach

For the test product to be bioequivalent to the reference standard, the 90% confidence interval for $\mu_T-\mu_R$ must be contained within the limits $\left[\frac{1}{m},\ m\right]$ in the original scale for each IVPT endpoint tested, where m = 1.2500.

APPENDIX I (EXAMPLE SAS CODE)

This section provides an example SAS code for use in determining BE in an in vitro permeation test (IVPT) study with either a balanced data set or an unbalanced data set. The example data sets, "Data-Balanced.csv" and "Data-Unbalanced.csv", are provided in Appendix II.

```
/*
INPUT
* dat = name of the data set
* don = column name of donor variable in dat
* reps = column name of replicate variable in dat
* trt = column name of treatment variable in dat (treatment variable: 'T',
'R')
* ly = column name of log-transformed endpoint in dat

OUTPUT: result
* Swr = estimated within-donor standard deviation of reference
* lpointest = point estimate for mean difference in log scale
* testmean = test mean estimate in original scale
* refmean = reference mean estimate in original scale
* pointest = point estimate transformed back to original scale
* ub = SABE 95% upper confidence bound
* (l, u) = ABE 90% CI for mean difference transformed back to original scale
```

```
%MACRO ivpt(dat=, don=, reps=, trt=, ly=);

      * Remove missing values before analysis;
      DATA wdat;
            SET &dat;
            IF &ly = . THEN DELETE;
      RUN;

      * Create the data sets for test & reference;
      DATA tdat;
        SET wdat;
        IF &trt = 'T';
      RUN;

      DATA rdat;
        SET wdat;
        if &trt = 'R';
      RUN;

      * Sort tdat and rdat by donor id and reps id;
      PROC SORT DATA=tdat;
        BY &don &reps;
      RUN;

      PROC SORT DATA=rdat;
        BY &don &reps;
      RUN;

      * Determine if the data is balanced or unbalanced;
      PROC SQL;
        CREATE TABLE num as
        SELECT &don, &trt, n(&don) as nrep
        FROM wdat
        GROUP BY &don, &trt;

        CREATE TABLE unum as
        SELECT DISTINCT(nrep) as nr
        FROM num;

        CREATE TABLE rcount as
        SELECT COUNT(*) as nnr
        FROM unum;

        DROP TABLE num, unum;
      QUIT;

      DATA _NULL_;
        SET rcount;
        CALL SYMPUT("nnr", nnr);
      RUN;

      %IF &nnr=1 %THEN %DO;   * if the data is balanced;

        * Calculate the necessary quantities;
        PROC SQL;
          CREATE TABLE tmp1 as
```

```
    FROM tdat GROUP BY &don;

    CREATE TABLE tmp2 as
    SELECT &don, mean(&ly) as mref, var(&ly) as vref, n(&ly) as rr
    FROM rdat GROUP BY &don;

    CREATE TABLE mergetmp0 as
    SELECT * FROM tmp1 as tmp1
    FULL JOIN tmp2 as tmp2
    on tmp1.&don = tmp2.&don;

    CREATE TABLE mergetmp as
    SELECT *, mtest-mref as Ij
    FROM mergetmp0;

    DROP TABLE tmp1, tmp2, mergetmp0;
QUIT;

PROC IML;
  USE mergetmp;
    READ ALL VAR {&don "vref" "rr" "Ij" "mtest" "mref"};
    m = 1.2500;
    alpha = 0.05;

    * Determine Swr;
    Swr2 = mean(vref);
    Swr = sqrt(Swr2);
    Ihat = mean(Ij);
    SI2 = var(Ij);
    nd = nrow(&don);
    nr = rr[1,1];
    df = (nr-1)*nd;

    * Treatment means;
    testmean = exp(mean(mtest));
    refmean = exp(mean(mref));

    * SABE for balanced data;
    theta = (log(m)/0.25)**2;
    qchi = quantile('chisq', 1-alpha, df);
    qt = quantile('t', 1-alpha, nd-1);
      =         - SI2/nd;
    y = - theta * Swr2;
    xp = ( abs(Ihat) + qt * sqrt(SI2/nd) )**2;
    yp = - theta * df * Swr2 / qchi;
    v = sign(xp-x) * (xp-x)**2 + sign(yp-y) * (yp-y)**2;
    ub = x + y + sign(v)*sqrt(abs(v));

    * ABE  for balanced data;
    se = sqrt(SI2/nd);
    logl = Ihat - qt * se;
    logu = Ihat + qt * se;

    l = exp(logl);
    u = exp(logu);

    * Rename the point estimate;
```

```
        lpointest = Ihat;
        pointest = exp(Ihat);

        CREATE result var {Swr lpointest
                                    testmean refmean pointest ub l u};
        APPEND;
        CLOSE result;
    QUIT;

    PROC PRINT DATA = result noobs;
      TITLE "IVPT Study Data Analysis Results for &ly: Balanced Data";
    RUN;

  %END;

  %ELSE %DO;   * if the data is unbalanced;

    * Estimate the mean difference;
    PROC MIXED DATA = wdat;
      CLASS &don &trt;
        MODEL &ly = &don &trt;
        ESTIMATE "&ly Test-Ref" &trt -1 1/cl alpha=0.1;

        ODS OUTPUT ESTIMATES = iout;
        ODS OUTPUT LSMEANS = mout;
    RUN; QUIT;

    * Calculate the necessary quantities;
    PROC SQL;
      CREATE TABLE tmp1 as
        SELECT &don, mean(&ly) as mref, n(&ly) as rr
        FROM rdat GROUP BY &don;

        CREATE TABLE tmp2 as
        SELECT count(*) as nd, sum(rr) as rstar
        FROM tmp1;
    QUIT;

    PROC IML;
      USE rdat;
        READ ALL VAR {&ly};

        USE tmp1;
        READ ALL VAR {&don "mref" "rr"};

        USE tmp2;
        READ ALL VAR {"nd" "rstar"};

        USE iout;
        READ ALL VAR {"estimate" "stderr" "df"};

        USE mout;
        READ ALL VAR {"estimate"} into lsmean;

        m = 1.2500;
        alpha = 0.05;
```

```
            * Determine Swr;
            mref2 = repeat(mref, rr);
            mref2 = shape(mref2, rstar, 1);

            Swr2 = sum( (&ly - mref2)##2 ) / (rstar - nd);
            Swr = sqrt(Swr2);

            * Treatment means;
            testmean = exp(lsmean[2,1]);
            refmean = exp(lsmean[1,1]);

            * SABE for unbalanced data;
            theta =  (log(m)/0.25)**2;
            qchi = quantile('chisq', 1-alpha, rstar-nd);
               =

            estimate = estimate[1,1];
            stderr = stderr[1,1];

            x = estimate**2 - stderr**2;
            y = - theta * Swr2;

            xp = (abs(estimate) + qt * stderr)** 2;
            yp = - theta * (rstar - nd) * Swr2 / qchi;

            v = sign(xp-x) * (xp-x)**2 + sign(yp-y) * (yp-y)**2;
            ub = x + y + sign(v)*sqrt(abs(v));

            * ABE for unbalanced data;
            logl = estimate - qt*stderr;
            logu = estimate + qt*stderr;

            l = exp(logl);
            u = exp(logu);

            * Rename the point estimate;
            lpointest = estimate;
            pointest = exp(estimate);

            CREATE result var {Swr lpointest
                                        testmean refmean pointest ub l u};
            APPEND;
            CLOSE result;
          QUIT;

          PROC PRINT DATA = result noobs;
            TITLE "IVPT Study Data Analysis Results for &ly: Unbalanced Data";
          RUN;

        %END;

%MEND ivpt;

proc import datafile = "Data-Balanced.csv"
      out = bdat
      dbms = csv
      replace;
```

```
run;

%ivpt(dat=bdat, don=donor, reps=replicate, trt=treat, ly=LAMT)


proc import datafile = "Data-Unbalanced.csv"

      dbms = csv
      replace;
run;

%ivpt(dat=udat, don=donor, reps=replicate, trt=treat, ly=LAMT)
```

APPENDIX II (NUMERICAL EXAMPLES)

This section provides numerical examples using simulated data sets illustrating the determination of BE.

(a) Balanced Data

The simulated data set "Data-Balanced.csv" shown below provides an example of in vitro permeation test (IVPT) study outcomes when the data is balanced. The SAS output and the determination of BE for LAMT in this data set follows.

Data-Balanced.csv

donor	replicate	treat	AMT	Jmax	LAMT	LJmax
1	1	T	2.361749	0.081326	0.859402	-2.50929
1	2	T	0.916571	0.041008	-0.08712	-3.19398
1	3	T	1.246243	0.038537	0.220133	-3.25613
1	4	T	0.890018	0.04296	-0.11651	-3.14747
1	5	T	0.663551	0.031219	-0.41015	-3.46674
1	6	T	0.479143	0.015747	-0.73576	-4.15108
2	1	T	0.998845	0.030073	-0.00116	-3.50412
2	2	T	0.814457	0.061644	-0.20523	-2.78637
2	3	T	0.648741	0.019984	-0.43272	-3.91283
2	4	T	1.142716	0.044332	0.133408	-3.11604
2	5	T	0.767291	0.028453	-0.26489	-3.55951
2	6	T	1.392406	0.166782	0.331033	-1.79107
3	1	T	1.388867	0.096822	0.328488	-2.33488
3	2	T	0.45757	0.02184	-0.78182	-3.82402
3	3	T	1.377438	0.0651	0.320226	-2.73182
3	4	T	0.870988	0.073199	-0.13813	-2.61457
3	5	T	1.753523	0.067281	0.561627	-2.69888
3	6	T	0.995674	0.116414	-0.00434	-2.15061

4	1	T	0.811458	0.053465	-0.20892	-2.92872
4	2	T	0.913538	0.060217	-0.09043	-2.8098
4	3	T	2.251438	0.083596	0.811569	-2.48176
4	4	T	1.163818	0.054213	0.151706	-2.91483
4	5	T	1.027813	0.065446	0.027433	-2.72653
4	6	T	1.081988	0.062279	0.078801	-2.77614
5	1	T	1.275517	0.069859	0.243352	-2.66128
5	2	T	1.231649	0.051342	0.208354	-2.96924
5	3	T	1.454325	0.161016	0.374542	-1.82625
5	4	T	1.195989	0.064734	0.178973	-2.73746
5	5	T	2.07678	0.088355	0.730819	-2.42639
5	6	T	1.893399	0.093223	0.638374	-2.37276
6	1	T	1.564164	0.137378	0.447352	-1.98502
6	2	T	1.504557	0.0728	0.408499	-2.62004
6	3	T	1.049724	0.064531	0.048527	-2.74061
6	4	T	1.047633	0.043859	0.046533	-3.12676
6	5	T	1.159634	0.09236	0.148105	-2.38206
6	6	T	1.129313	0.06546	0.12161	-2.72632
1	1	R	1.598636	0.04239	0.469151	-3.16084
1	2	R	2.24476	0.117486	0.808599	-2.14143
1	3	R	1.60912	0.044199	0.475687	-3.11906
1	4	R	1.8834	0.066452	0.633079	-2.71127
1	5	R	1.101948	0.031705	0.097079	-3.45129
1	6	R	1.165342	0.034002	0.153015	-3.38133
2	1	R	0.622369	0.052794	-0.47422	-2.94135
2	2	R	0.833337	0.033419	-0.18232	-3.39863
2	3	R	0.386763	0.029507	-0.94994	-3.52311
2	4	R	0.294178	0.02005	-1.22357	-3.9095
2	5	R	0.851759	0.03968	-0.16045	-3.2269
2	6	R	0.677715	0.032379	-0.38903	-3.43024
3	1	R	0.96461	0.042626	-0.03603	-3.15528
3	2	R	0.838346	0.101628	-0.17632	-2.28643
3	3	R	0.130884	0.008774	-2.03344	-4.73601
3	4	R	0.635926	0.039118	-0.45267	-3.24118
3	5	R	0.804131	0.114582	-0.21799	-2.16646
3	6	R	2.324877	0.229704	0.843667	-1.47096
4	1	R	1.694799	0.088825	0.527564	-2.42109
4	2	R	0.977661	0.030392	-0.02259	-3.49356
4	3	R	3.13529	0.217896	1.142722	-1.52374
4	4	R	0.922805	0.040161	-0.08034	-3.21485
4	5	R	1.504834	0.082443	0.408683	-2.49565

4	6	R	1.330167	0.055237	0.285305	-2.89612
5	1	R	2.104036	0.101673	0.743858	-2.28599
5	2	R	0.842231	0.094771	-0.1717	-2.35629
5	3	R	0.985656	0.081963	-0.01445	-2.50148
5	4	R	0.931461	0.069496	-0.071	-2.66648
5	5	R	1.580578	0.059193	0.45779	-2.82695
5	6	R	1.209059	0.067989	0.189842	-2.68841
6	1	R	1.038591	0.037859	0.037865	-3.27389
6	2	R	1.064539	0.049079	0.062542	-3.01433
6	3	R	0.795337	0.028705	-0.22899	-3.55068
6	4	R	0.922567	0.035194	-0.0806	-3.34689
6	5	R	0.780047	0.034144	-0.2484	-3.37716
6	6	R	1.415222	0.066506	0.347286	-2.71046

SAS Output

SWR	LPOINTEST	TESTMEAN	REFMEAN	POINTEST	UB	L	U
0.50242	0.096445	1.11571	1.01313	1.10125	-0.022242	0.80470	1.50708

The estimated within-donor standard deviation of the RS is 0.5024, which is greater than 0.294. Using the SABE approach, the 95% upper confidence bound is -0.0222 < 0 and the point estimate of 1.1013 is within the BE limit of [0.8000, 1.2500]. Thus, BE for AMT can be concluded. The BE test for J_{max} can be performed similarly.

(b) Unbalanced Data

The simulated data set "Data-Unbalanced.csv" shown below provides an example of IVPT study outcomes when the data is unbalanced. The SAS output and the determination of BE for LAMT in this data set follows.

Data-Unbalanced.csv

donor	replicate	treat	AMT	Jmax	LAMT	LJmax
1	1	T	2.361749	0.081326	0.859402	-2.50929
1	2	T	0.916571	0.041008	-0.08712	-3.19398
1	3	T	1.246243	0.038537	0.220133	-3.25613
1	4	T	0.890018	0.04296	-0.11651	-3.14747
1	5	T	0.663551	0.031219	-0.41015	-3.46674
1	6	T	0.479143	0.015747	-0.73576	-4.15108
2	1	T	0.998845	0.030073	-0.00116	-3.50412
2	2	T	0.814457	0.061644	-0.20523	-2.78637
2	3	T	0.648741	0.019984	-0.43272	-3.91283

2	4	T	0.767291	0.028453	-0.26489	-3.55951
2	5	T	1.392406	0.166782	0.331033	-1.79107
3	1	T	0.45757	0.02184	-0.78182	-3.82402
3	2	T	1.377438	0.0651	0.320226	-2.73182
3	3	T	2.170988	0.073199	0.775183	-2.61457
3	4	T	1.753523	0.067281	0.561627	-2.69888
3	5	T	0.995674	0.116414	-0.00434	-2.15061
4	1	T	0.811458	0.053465	-0.20892	-2.92872
4	2	T	0.913538	0.060217	-0.09043	-2.8098
4	3	T	1.251438	0.083596	0.224293	-2.48176
4	4	T	1.163818	0.054213	0.151706	-2.91483
4	5	T	1.027813	0.065446	0.027433	-2.72653
4	6	T	1.081988	0.062279	0.078801	-2.77614
5	1	T	1.275517	0.069859	0.243352	-2.66128
5	2	T	1.231649	0.051342	0.208354	-2.96924
5	3	T	2.454324	0.161016	0.897852	-1.82625
5	4	T	1.195989	0.064734	0.178973	-2.73746
5	5	T	2.07678	0.088355	0.730819	-2.42639
5	6	T	1.893399	0.093223	0.638374	-2.37276
6	1	T	1.564164	0.137378	0.447352	-1.98502
6	2	T	1.049724	0.064531	0.048527	-2.74061
6	3	T	1.047633	0.043859	0.046533	-3.12676
6	4	T	1.159634	0.09236	0.148105	-2.38206
1	1	R	1.598636	0.04239	0.469151	-3.16084
1	2	R	2.24476	0.117486	0.808599	-2.14143
1	3	R	1.60912	0.044199	0.475687	-3.11906
1	4	R	1.8834	0.066452	0.633079	-2.71127
1	5	R	1.101948	0.031705	0.097079	-3.45129
1	6	R	1.165342	0.034002	0.153015	-3.38133
2	1	R	0.622369	0.052794	-0.47422	-2.94135
2	2	R	0.833337	0.033419	-0.18232	-3.39863
2	3	R	0.386763	0.029507	-0.94994	-3.52311
2	4	R	0.851759	0.03968	-0.16045	-3.2269
2	5	R	0.677715	0.032379	-0.38903	-3.43024
3	1	R	0.838346	0.101628	-0.17632	-2.28643
3	2	R	0.130884	0.008774	-2.03344	-4.73601
3	3	R	0.635926	0.039118	-0.45267	-3.24118
3	4	R	0.804131	0.114582	-0.21799	-2.16646
3	5	R	2.324877	0.229704	0.843667	-1.47096
4	1	R	1.694799	0.088825	0.527564	-2.42109
4	2	R	0.977661	0.030392	-0.02259	-3.49356

4	3	R	3.13529	0.217896	1.142722	-1.52374
4	4	R	0.922805	0.040161	-0.08034	-3.21485
4	5	R	1.504834	0.082443	0.408683	-2.49565
4	6	R	1.330167	0.055237	0.285305	-2.89612
5	1	R	2.104036	0.101673	0.743858	-2.28599
5	2	R	0.842231	0.094771	-0.1717	-2.35629
5	3	R	0.985656	0.081963	-0.01445	-2.50148
5	4	R	0.931461	0.069496	-0.071	-2.66648
5	5	R	1.580578	0.059193	0.45779	-2.82695
5	6	R	1.209059	0.067989	0.189842	-2.68841
6	1	R	1.038591	0.037859	0.037865	-3.27389
6	2	R	0.795337	0.028705	-0.22899	-3.55068
6	3	R	0.922567	0.035194	-0.0806	-3.34689
6	4	R	0.780047	0.034144	-0.2484	-3.37716
6	5	R	1.415222	0.066506	0.347286	-2.71046

SAS Output

SWR	LPOINTEST	TESTMEAN	REFMEAN	POINTEST	UB	L	U
0.50651	0.067494	1.10723	1.03497	1.06982	-0.10907	0.87627	1.30613

The estimated within-donor standard deviation of the RS is 0.5065, which is greater than 0.294. Using the SABE approach, the 95% upper confidence bound is -0.1091 < 0 and the point estimate of 1.0698 is within the BE limit of [0.8000, 1.2500]. Thus, BE for AMT can be concluded. The BE test for J_{max} can be performed similarly.

APPENDIX III (EXAMPLE R CODE)

This section provides an example of R code that performs the same calculations as the SAS code in Appendix I.

```
## INPUT

# DAT = a data frame
# DON = column name of donor variable in DAT (donor variable: numeric)
# REPS = column name of replicate variable in DAT
#        (replicate variable: numeric)

#        (treatment variable: "T", "R")
# LY = column name of log-transformed endpoint in DAT


## OUTPUT
```

```
# BU = balanced data or unbalanced data
# Swr = estimated within-donor standard deviation of reference
# Ihat = point estimate for mean difference in log scale
# testMean = test mean estimate in original scale
# refMean = reference mean estimate in original scale
# pointest = point estimate transformed back to original scale
# UB = SABE 95% upper confidence bound
# CI = ABE 90% CI for mean difference transformed back to original scale

ivpt <- function(DAT, DON, REPS, TRT, LY){

  # Remove missing values before analysis
  DAT <- DAT[!is.na(DAT[[LY]]),]

  # If zero values in AMT or Jmax are not imputed,
  # remove them to avoid a computational burden
  DAT <- DAT[!(DAT[[LY]]==-Inf),]

  # Create the data sets for test & reference
  tdat <- DAT[DAT[[TRT]]=="T",]
  rdat <- DAT[DAT[[TRT]]=="R",]

  # Sort tdat and rdat by don and reps
  ii1 <- order(tdat[[DON]], tdat[[REPS]])
  tdat <- tdat[ii1,]
  ii2 <- order(rdat[[DON]], rdat[[REPS]])
  rdat <- rdat[ii2,]

  # Determine the numbers of replicates from each donor for T & R
  rT <- as.vector(table(tdat[[DON]]))
  rR <- as.vector(table(rdat[[DON]]))

  nd <- length(unique(tdat[[DON]])) # the number of donors
  rstar <- sum(rR)        # the total number of replicates in R group


  # Set m and alpha
  m <- 1.2500
  alpha <- 0.05
  theta <- (log(m)/0.25)^2

  if(length(nr)==1){   # if the data is balanced

    BU <- "Balanced"

    # Determine SWR
    mRef <- tapply(rdat[[LY]], rdat[[DON]], mean, na.rm=TRUE)
    vv <- tapply(rdat[[LY]], rdat[[DON]], var, na.rm=TRUE)
    Swr2 <- sum(vv) / nd
    Swr <- sqrt(Swr2)

    mTest <- tapply(tdat[[LY]], tdat[[DON]], mean, na.rm=TRUE)
    Ij <- mTest - mRef

    SI2 <- var(Ij, na.rm=TRUE)
```

```
  # Treatment means
  testMean <- exp(mean(mTest))
  refMean <- exp(mean(mRef))

  # SABE for balanced data
  X <- Ihat^2 - SI2 / nd
  Y <- - theta * Swr2
  Xp <- ( abs(Ihat) + qt(1-alpha, nd-1) * sqrt(SI2/nd) )^2
  Yp <- - theta*(nr-1)*nd*Swr2 / qchisq(1-alpha, (nr-1)*nd)
  V <- sign(Xp-X)*(Xp-X)^2 + sign(Yp-Y)*(Yp-Y)^2
  UB <- X + Y + sign(V)*sqrt(abs(V))

  # ABE for balanced data
  se <- sqrt(SI2/nd)
  L <- Ihat - qt(1-alpha, nd-1)*se
  U <- Ihat + qt(1-alpha, nd-1)*se

}else{              # if the data is unbalanced

  BU <- "Unbalanced"

  # Determine SWR
  mRef <- tapply(rdat[[LY]], rdat[[DON]], mean, na.rm=TRUE)
  vv <- sum( (rdat[[LY]] - rep(mRef, times=rR))^2 )
  Swr2 <- vv / (rstar - nd)
  Swr <- sqrt(Swr2)

  # Estimate the mean difference
  DAT[[DON]] <- factor(DAT[[DON]])
  f <- as.formula(paste(LY, "~", DON, "+", TRT))
  obj <- lm(f, data = DAT)
  tname <- paste0(TRT, "T")
  tcoef <- summary(obj)$coef[rownames(summary(obj)$coef)==tname,]
  Ihat <- as.numeric(tcoef[1])
  se <- as.numeric(tcoef[2])
  dfstar <- summary(obj)$df[2]

  # Treatment means
  udon <- unique(DAT[[DON]])
  newdat1 <- data.frame(DON = udon, TRT = rep("T", length(udon)))
  newdat2 <- data.frame(DON = udon, TRT = rep("R", length(udon)))
  colnames(newdat1) <- c(DON, TRT)
  colnames(newdat2) <- c(DON, TRT)
  testMean <- exp(mean(predict(obj, newdata = newdat1)))
  refMean <- exp(mean(predict(obj, newdata = newdat2)))

  # SABE for unbalanced data
  X <- Ihat^2 - se^2
  Y <- - theta * Swr2
  Xp <- ( abs(Ihat) + qt(1-alpha, dfstar)*se )^2
  Yp <- - theta*(rstar-nd)*Swr2 / qchisq(1-alpha, rstar-nd)
  V <- sign(Xp-X)*(Xp-X)^2 + sign(Yp-Y)*(Yp-Y)^2
  UB <- X + Y + sign(V)*sqrt(abs(V))

  # ABE for unbalanced data
  L <- Ihat - qt(1-alpha, dfstar)*se
```

```
  }

  result <- list(BU=BU, Swr=Swr, Ihat=Ihat,
                 testMean=testMean, refMean=refMean, pointest=exp(Ihat),
                 UB=UB, CI=exp(c(L, U)))

  return(result)
}

bdat <- read.csv("Data-Balanced.csv")


udat <- read.csv("Data-Unbalanced.csv")
res2 <- ivpt(udat, "donor", "replicate", "treat", "LAMT")
```